

## DOCUMENT RESUME

ED 127 349

TM 005 439

AUTHOR Judd, Charles M.; Steil, Janice Ingham  
TITLE The Evaluation of Affective Goals in Two High School Social Studies Curricula.  
PUB DATE [Apr 76]  
NOTE 20p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Francisco, California, April 19-23, 1976)  
EDRS PRICE MF-\$0.83 HC-\$1.67 Plus Postage.  
DESCRIPTORS \*Affective Objectives; \*Changing Attitudes; Dogmatism; \*Evaluation; Evaluation Methods; High School Curriculum; High School Students; \*Humanistic Education; Multiple Regression Analysis; Secondary Education; Sex Differences; Social Studies; Student Opinion  
IDENTIFIERS Rokeach (Milton)

## ABSTRACT

A study was conducted to assess certain affective outcomes in two high school social studies curricula. A theory of affective growth based on Rokeach (1960) was developed and tested in a quasi-experimental pre-post design. Data were analyzed using Multiple Regression Analysis to control for pre-course responses. Sex, curriculum, a motivation/ability proxy, and two interaction terms were entered as independent variables. Significant differences in outcomes were found by type of curriculum and by sex, controlling for pre-course responses. This study affirms the importance of, and provides a methodology for, the assessment of affective educational outcomes. (Author/BW) were analyzed

\*\*\*\*\*  
\* Documents acquired by ERIC include many informal unpublished \*  
\* materials not available from other sources. ERIC makes every effort \*  
\* to obtain the best copy available. Nevertheless, items of marginal \*  
\* reproducibility are often encountered and this affects the quality \*  
\* of the microfiche and hardcopy reproductions ERIC makes available \*  
\* via the ERIC Document Reproduction Service (EDRS). EDRS is not \*  
\* responsible for the quality of the original document. Reproductions \*  
\* supplied by EDRS are the best that can be made from the original. \*  
\*\*\*\*\*

ED127349

The Evaluation of Affective Goals  
in Two High School Social Studies Curricula

Charles M. Judd & Janice Ingham Steil  
Teachers College, Columbia University

Paper presented at annual meetings,  
American Educational Research Association,  
San Francisco, April, 1976.

U.S. DEPARTMENT OF HEALTH  
EDUCATION & WELFARE  
NATIONAL INSTITUTE OF  
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY.

TM005 439

## I. INTRODUCTION

This project was initiated in response to a teacher's request for help in evaluating the "affective outcomes" of a high school social studies curriculum. After several meetings focusing on task clarification, the teacher asked if it would be possible to measure anticipated student changes in degree of tolerance for others' beliefs as a result of their participation in a one semester Asian studies course.

In many ways, this teacher's request, as initially posed, and even subsequent to numerous discussion, reflects several characteristics of the broader educational community from which it emerged. An intuitive belief exists on the part of educators that school curriculum has affective as well as cognitive outcomes. The two, however, are often viewed not only differentially but dichotomously.

Cognitive outcomes have traditionally been viewed as more legitimately the task of the school and, perhaps consequently, have been more effectively translated into explicit behavioral objectives in accordance with existing psychological and educational theory. Likewise, progress toward attaining cognitive objectives is more explicitly and consistently evaluated in terms of an extensive repertoire of measurement techniques.

Expected affective outcomes, on the other hand, are often more implicit than explicit; and even when explicitly stated are seldom derived from coherent educational or psychological theory; defined in behavioral terms; or subjected to any sort of systematic evaluation.

It should be pointed out that social scientists have done little to alter this approach. Other than Kolhberg and Piaget's work in moral development, there is little coherent theory from which educators can draw in their work in the affective realm. The literature contains few if any instances of thoughtful attempts to measure affective outcomes of educational curricula and the classic, large scale input-output studies (Coleman, et al., 1966) have also concentrated primarily on cognitive measures.

In meeting the challenge of these issues, the authors have drawn heavily from the work of Milton Rokeach. His work seemed particularly relevant for several reasons. First, he presents a coherent and fairly comprehensive theory of the organization of belief systems which is concerned with structure rather than content. Second, he presents a cognitive theory of affective system change. Third, he specifically relates his ideas to the concept of open versus closed mindedness: a concept which specifically subsumes tolerance versus intolerance. Fourth, he has done preliminary work on the development of instruments for the purpose of measuring extent of open versus closed-mindedness.

Rokeach conceives of all cognitive systems as being organized into two interdependent parts, a belief system (which represents the total universe of a person's beliefs about the physical world, the social world and the self) and a disbelief system. It is assumed that not all beliefs are equally important to the individual. Beliefs vary along a central-peripheral dimension which ranges from type A (primitive beliefs reinforced by a unanimous social consensus among all of one's reference persons and groups) to

type E (inconsequential beliefs relating to arbitrary matters of taste).

Further, the disbelief system is conceived to be composed of several subsystems arranged along a continuum of similarity to the belief system and it is hypothesized that disbelief subsystems that are similar to the belief system are more acceptable than less similar ones.

Rokeach uses the terms closed or dogmatic interchangeably.

To say that a person is dogmatic or that his belief system is closed is to say something about the way he believes and the way he thinks--not only about single issues but about networks of issues.

Ideological dogmatism according to Rokeach refers to: a) a closed way of thinking which can be associated with any ideology regardless of content; b) an authoritarian outlook on life; c) intolerance of those with opposing beliefs; d) suffrance of those with similar beliefs (1960).

Rokeach asserts that a closed belief system will lead to a stronger acceptance of the belief system but also to a stronger rejection of all disbelief subsystems. It should lead the person to have a greater discrepancy between what he knows about the belief system and what he knows about the disbelief system. It should also lead to less differentiation within the disbelief system as a whole and to an increasing inability to distinguish among disbelief subsystems the more dissimilar they are from the belief system.

From this, he asserts that the defining characteristics of openness--closedness are: degree of rejection of disbelief subsystems, degree of differentiation of beliefs system as compared with the disbelief system, and the degree of differentiation within the

disbelief system. Further, Rokeach hypothesizes that the more closed a person's belief system the more will others be evaluated according to the authorities they line up with; the more will others be evaluated according to their agreement or disagreement with the person's own system; and the greater will be the difficulty for the person to discriminate between and separately evaluate a belief and the person holding the belief.

On the basis of this theoretical background, the authors derived four hypotheses specifying predicted changes in student behavior as a result of participation in the Asian Studies course. A measurement instrument asking students to respond to questions relating to single and paired statements from the student's disbelief system was developed to test the predictions.

Hypothesis I: predicts a) an increased tendency not to reject the source of opinions or attitudes while nevertheless rejecting the opinions themselves; and b) an increase in the level of uncertainty in those instances in which the source does continue to be rejected with the opinion. In Rokeach's (1960) terminology, (a) predicts a decreased tendency to make "opinionated statements."

Hypothesis II: predicts an increased tendency to differentiate among opinions, beliefs, or life styles which are not one's own or with which one disagrees. In Rokeach's (1960) terminology: an increased tendency towards "differentiation within the disbelief system." (p. 38)

Hypothesis III: predicts an increased tendency to perceive similarities between one's own value system and way of life and differing value systems and ways of life.



Hypothesis IV: predicts an increased tendency to perceive opinions and values as functional and at least partially attributable to external causes (i.e. of cultural and social causes) as opposed to only internal causes (i.e. result of personality).

From Rokeach's position that highly dogmatic persons tend to greater rejection of the source of a statement with which they disagree than lower dogmatic persons, it seems reasonable to hypothesize that as students become less opinionated and more tolerant as a result of the Asian Studies experience, there will be a decrease in their tendency to explain behaviors (i.e. as represented by the opinion statements) to only internal causes and a concomitant increase in their tendency to at least partially attribute them to external (sociocultural) factors. Thus, it is hypothesized that if a student is presented with any given value, opinion statement, or belief, and he is asked why that particular belief is held, or what it is due to, he will increasingly tend, as a result of the course, to attribute the belief to social and cultural causes exterior to the source of the belief.

## Method

### Overview of Method

At the end of the school semester, a questionnaire entitled "Student Opinion Questionnaire" was administered to 99 students in one of five social studies classes taught by the same teacher.

Of the five classes, two had followed a teacher developed Asian Studies curriculum, while the other three followed a curriculum concentrating on American law and jurisprudence. In addition, half of the students had completed an identical questionnaire at the start of the semester. The researchers were introduced to the class as doctoral students from Teachers College, Columbia University, who were interested in "opinions" of high school students. The reason for this was to disassociate the administration of the questionnaire from the students' school related studies, and particularly from the social studies courses.

The questionnaire was developed by the researchers without any knowledge of the specific content of the two curricula. In addition, the teacher was not informed of the study's hypotheses nor of the content of the questionnaire until after the end of the semester. These precautions were taken to insure that the students were not explicitly taught how to answer the questionnaire in a way which would support the hypotheses.

#### Research Design

The research design employed in this study was quasi-experimental in nature, due to the fact that students were not assigned to the "treatments" (i.e. the two curricula) randomly, but rather chose which course they wanted to take. Because of this fact and the potential dangers of resulting selection biases, it was necessary to include a pre-test in the research design, so that self selection differences, prior to the two treatments, could be statistically controlled for in the analysis of post-test responses.



However, giving a pre-test to subjects frequently induces a "testing" artifact (Campbell & Stanley, 1965) or even a testing by treatment interaction artifact, which might also have confounded any post-test differences between treatments. Therefore, each treatment group was randomly divided in half, one half of each group then received the test instrument at the beginning of the semester; the other half of each group had a study hall during that time. Those who had been pre-tested were asked not to disclose the nature of the instrument to those who were not so tested. At the end of the semester, all students in both treatments were given the testing instrument.

In addition to the potential problems due to self-selection and a potential testing artifact, it should be apparent that neither treatment is a true control for the other. For instance, the Law curriculum is much more than just the absence of the Asian Studies curriculum. Therefore it is essential in the analysis of the post-test data that differences between treatments not be exclusively interpreted as resulting from only one curriculum, but that such differences may in part be due to either or both curricula. It would have been somewhat easier had we only dealt with one treatment group, for instance the Asian Studies students, and analyzed "change scores" between pre-test and post-test responses. Our inclusion of a "control" treatment, even though not a true control, was necessitated by the danger of a regression artifact when dealing with change scores (Cronbach & Furby, 1970).

### Measures

In order to assess hypotheses 1, 3, and 4, students were

presented with four opinion statements with which it was judged they would disagree. All four statements were beliefs about cultural practices, differing from most of those held by the students and their families. The student was asked a number of questions about each of these opinion statements. First they were asked to indicate on a 7-point Likert scale the extent to which they agreed or disagreed with the opinion. Responses to this question were used as controls in the testing of all hypotheses. Next, to test Hypothesis 1, students were asked to indicate the extent to which they thought they "would like as a friend a person who believed" the statement. The third questions, also a test of Hypothesis 1, asked them to rate the extent to which they were "certain" of their answer to the previous question.

The fourth and fifth questions, on each of the four opinion statements, operationalized Hypothesis 3. The students were asked how much they thought they "would agree or disagree with other opinions of someone who believed" the statement. They were also asked how certain they were of their answer to this question. Finally, to test the fourth hypothesis, students were asked to write in a sentence or two "why (they) thought someone would believe" the statement. In total then, students were asked five Likert scale and one open-ended question on each of the four opinion statements.

To test the second hypothesis, four pairs of opinion statements were presented to the students and a number of questions asked about each of the pairs. While the statements were chosen because they represented different cultural beliefs than those held by most

American high school students, any two statements which made up a pair were irrelevant to and independent of each other. For each pair the student was asked to rate the extent to which he or she agreed with each of the statements in the pair. Responses to these two questions were controlled for in the analysis of responses of the third question which asked the student: "If these two statements were made by two different people whom you just met, how much do you think those two people would agree or disagree on other things?" The fourth question asked the student to indicate how certain they were of their answer to the previous question.

In addition to these tests of the hypotheses, students were asked to rank order six of Rokeach's Terminal Values (1973) in terms of how important those values were to them personally.

In addition to these questionnaire items gathered from the students, the following data were also available on each student: sex, year in high school, cumulative grade point average, and tracking level (all students had been assigned to one of three tracks).

#### Methods of Analysis

In order to examine the data relevant to our fourth hypothesis, students' open-ended responses to questions about why the beliefs might be held were rated by judges, blind both to condition and to the hypotheses of the study, on a variety of scales. Rating were made on the extent to which the student answered the question, the nature of the attribution made by the student, the apparent confidence

in the attribution, and the degree of evaluation of both statement and of the person making the statement inherent in the response. Three judges independently rated all responses on the eight scales with a mean inter-judge across scale reliability coefficient of .745.

All responses to the questionnaires administered to all students at the end of the semester were analyzed using Multiple Regression Analysis (Kerlinger & Pedhauzer, 1974). The following independent variables were entered in all of the regression equations:

a. Responses to the same test item at the beginning of the semester. This was included to control for self-selection factors in the two treatment groups. For those students who had not taken the pre-test, the mean score for those in the same treatment group who had taken it was assigned.

b. A dummy variable, coded 1 if the student had taken the pre-test, coded 0 otherwise. This independent variable was included to control for any possible "testing" artifact.

c. An interaction term between treatment condition and whether or not the student had taken the pre-test. As explained previously, this interaction was included because it is conceivable that a "testing" artifact might have been brought out in one treatment group but not in the other, due to the content of questions and curricula.

d. The tracking level to which the student had been assigned. This measure was assumed to be a proxy for both the student's ability and motivation.

e. A sex dummy variable, coded 1 if the student was female, coded 0 if the student was male.

f. A treatment dummy variable, coded 1 if the student was in an Asian Studies class, coded 0 if the student was in a Law class.

g. An interaction term between treatment condition and the tracking level of the student.

h. An interaction term between treatment condition and sex of the student. This interaction was included because of the possibility that there might be treatment effects for one sex but not for the other.

Depending upon the specific dependent measure regressed, other independent variables were also included. For instance, the extent to which the student agreed or disagreed with the appropriate statement or statements was entered as an independent variable in all subsequent questions about that statement or statements.

### RESULTS

The complexity of our measures, testing each of our four hypotheses with eight or more different measures, argues against a detailed reporting of results in the present summary paper. We also feel that the major value of this study lies not in the specific findings, but rather in the theoretical and methodological suggestions for the evaluation of affective schooling outcomes. We therefore will only give summaries of results for each of our

hypotheses.

Hypothesis 1. Three sorts of measures were analyzed to test Hypothesis 1: First, students were asked to indicate the extent to which they thought "they would like as a friend" someone who believed each of the four statements, with which the students disagreed. Secondly, on the four questions which asked "Why would someone believe this statement?", judges rated the extent to which the answers were evaluative both of the person who believed the statement and of the statement itself. The former scores were analyzed to test Hypothesis 1. Finally, the third question asked students to rate how certain they were that they would like or dislike someone who believed each of the four statements. These uncertainty scores were also used to test the hypothesis. In all these analyses, the extent to which the student agreed or disagreed with the appropriate statement was controlled for.

Support for the hypothesis emerged from two of these three sets of measures. First, the judged evaluation of persons holding the beliefs showed effects on some statements for type of curriculum and its interaction with sex. Secondly the measure of certainty in liking or disliking such a person showed effects on some statements for type of curricula and for sex and their interaction. It seems that on these two measures, the Asian curriculum, as compared with the Law, affected students' responses in the predicted direction, and the certainty measure showed changes especially for female students, both across curricula and within the Asian courses.

Hypothesis 2. The second hypothesis was tested by asking questions about pairs of statements which constituted part of the



students' disbelief systems. Students were asked, first of all, to indicate the degree to which they agreed or disagreed with each statement of the four pairs. Subsequently for each pair, students were asked to indicate the extent to which they thought two people who believed the statements would agree or disagree about other beliefs. They were then asked to indicate the degree to which they were certain of their answer to the previous questions. Both of these last two responses were used as measures to test the second hypothesis.

For the first measure, there emerged a significant curriculum by sex interaction term for one of the pairs of statements, supportive of the hypothesis for female students. Females in the Asian Studies curriculum came to differentiate more between statements which constituted part of their disbelief systems. For the second measure, degree of certainty, females in both curricula became less certain of their responses when precourse levels of certainty were controlled for.

Hypothesis 3. The third hypothesis dealing with the degree of perceived similarities between aspects of the belief and disbelief systems was assessed by determining the extent to which students thought someone with whom they disagreed on one issue would also disagree with them on other issues. Students were also asked to indicate their certainty in their response to this question. None of the eight questions which were used to test this hypothesis showed any significant effects for type of curriculum or its interaction with sex of student. Thus the results failed to support affective change as a result of the course experience on this

hypothesis.

Hypothesis 4. Students were asked to indicate in one or two sentences why they thought people would believe each of the statements (excluding those which were paired) which constituted part of the students' disbelief systems. These responses were coded by trained judges in three ways which are relevant to this fourth hypothesis. First, judges rated the extent to which students attributed the statements to "personal, internal characteristics." Next, judges rated the extent to which attributions were made to "personal experience, family, etc."; and finally, judges rated the extent to which attributions were to "social and cultural norms."

Analysis revealed support for this hypothesis on questions concerning external attributions to personal experience. For one of the statements, students in the Asian Studies courses came increasingly to make such external attributions, and for another statement, females in the Asian Studies courses did so. In addition, for one of the statements, the Asian Studies curriculum led to significantly more external attributions to social and cultural norms, consistent with our hypothesis.

Finally, there were also treatment differences in the ranking of some of the terminal values, when pre-treatment ranks were controlled for. Thus, students in the Asian Studies courses came to view the value "Obedience" as less important than students in the Law courses. In addition, the value "Cleanliness" came to be ranked more highly by females in both curricula.

### DISCUSSION

Overall there seem to be a variety of results supportive of some of our hypotheses, on some of the statements. These affective effects seem particularly to be found among female students who were in the Asian Studies courses. There obviously emerge a number of questions concerning these results. First of all, why were some hypotheses demonstrated and not others? Next, why were the hypotheses demonstrated by some of the statements which were part of the students' disbelief system and not by others? And finally, why do our effects seem to be stronger for female students? We turn very briefly to these three questions.

The strongest affective changes in students' responses were found for the first and the fourth hypotheses. Thus, students came to make opinionated statements with less certainty and they made more external attributions as a result of the course. Support for the second hypothesis, predicting increased differentiation within the disbelief system, was only weak, and there was no support for the third hypothesis. There are at least two possible explanations for this difference in supportive results by hypotheses. The first is that our first hypothesis, as tested by the certainty questions, and the fourth deal with behaviors more susceptible to change than the second and third hypotheses. It could be argued that such a hierarchy makes some theoretical sense. The second possible explanation is that measures of each hypothesis differed in their reliability and validity. Unfortunately this latter explanation cannot be ruled out. Extensive testing of the validity

and ~~reliability~~ of our measures should be conducted prior to further use.

There are a variety of reasons why our effects were shown on some statements and not on others. First of all, there is evidence that affective changes are best shown when dealing with statements that are relatively less disagreed with. While there is evidence that all of our statements constituted part of the students' disbelief systems, most of our effects did occur on questions about statements which were less extremely disagreed with. Further, there is some evidence that topics or issues covered by the statements were differentially familiar and relevant to high school students. Thus statements dealing with differing practices surrounding courtship and marriage may be judged to be more important for high school students, and thus less apt to show response changes, than statements concerning bathing practices.

Turning to the issue that our effects seemed to be stronger among female students than among males, one ready explanation for this difference is the fact that the teacher was herself female. In a post semester interview, the teacher indicated that she had felt she had communicated more to her female students, and had at times felt that her authority had been challenged by the male students. This difference by sex of students suggests the need for further studies where the interaction between sex of teacher and sex of student can be examined in more detail.

While the results of this study are only moderately successful in terms of demonstrating our hypotheses, it is to be hoped that

the study will provide an impetus for theoretical refinement and improved research in the evaluation of affective schooling outcomes. The methodology employed, the use of multiple regression techniques within a quasi-experimental research design, is felt to have a potentially wide applicability; and the model of colleagueal collaboration could help bridge the gap that has traditionally existed between findings in and techniques of social science research and educational application and implementation.

### Bibliography

Campbell, D., & Stanley, D. Experimental and Quasi-Experimental Designs for Research. New York: Rand-McNally, 1963.

Coleman, J. et al. Equality of Educational Opportunity. Office of Education, Department of Health, Education, and Welfare, 1966.

Cronbach, L. J., & Furby, L. How should we measure change - or should we? Psychological Bulletin, 1970, 74, 68-80.

Kerlinger, F. N., & Pedhazur, E. J. Multiple Regression in Behavioral Research. New York: Holt, Rinehart & Winston, 1974.

Rokeach, M. The Open and Closed Mind. New York: Basic Books, 1960.

Rokeach, M. Beliefs, Attitudes and Values: A Theory of Organization and Change. Jossey Bass Inc., Pub., 1972.

Rokeach, M. The Nature of Human Values. New York: The Free Press, 1973.